Tc MRKSVCPKQKFFFSAFPFFFFFCVFPLISRTGQEKLLFDQKYKIIKGEKKEKKKNQRANRREHQQKREIMRFKKS   75

Tc FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIEEGADLGMVF   150

Tc MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVVLDTPAGLVRGTAHLQSGTESEVSNASIINVPSFLYQ   225

Tc QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC   300

Tc VEIYGPPTNPEANYKNVVIFGNRQADR [SPCGT] GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE   371

Tc   RIPGVKVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*   423

86

```
Tc                              RTGQEKLLFDQKYKIIKGEKKEKKKNQRANRREHQQKREIMRFKKS              75

Tc    FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIEEGADLGMVF    150

Tc    MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVVLDTPAGLVRGTAHLQSGTESEVSNASIINVPSFLYQ    225

Tc    QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC    300

Tc    VEIYGPPTNPEANYKNVVIFGNRQADR  SPCGT  GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE    371

Tc      RIPGVKVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*          423
```

Tc   MRKSVCPKQKFFFSAFPFFFFFCVFPLIS

```
Cs   --------------------------------------------------------------MKFSKG              6

Cs   IHAIDSHTMGEPTRIVVGGIPQINGETMADKKKYLEDNLDYVRTALMHEPRGHNDMFGSIITSSNNKEADFGIIF         81

Cs   MDGGGYLNMCGHGSIGAATVAVETGMVEMVEPVTNIN--MEAPAGLIKAKVMVEN---EKVKEVSITNVPSFLYM        151

Cs   EDAKLEVPSLNKTITFDISFGGSFFAIIHAKELGVKVETSQVDVLKKLGIEIRDLINEKIKVQHPELEHIKTVDL        226

Cs   VEIYDEPSNPEATYKNVVIFGQGQVDR │ SPCGT │ GTSAKLATLYKKGHLKIDEKEVYESITGTMFKGRV--LEET  297

Cs   KVGEFD-----------AIIPEITGGAYITGENHEVIDPEDPLKYGFTV*--  335
```

```
Pa   --------------------------------------------------------------------------MQR      3

Pa   IRIIDSHTGGEPTRLVIGGFPDLGQGDMAERRRLLGERHDAWRAACILEPRGSDVLVGALLCAPVDPEACAGVIF     73

Pa   FNNSGYLGMCGHGTIGLVASLAHLGRIGPGV------HRIETPVGEVEATLH-------EDGSVSVRNVPAYRYR    140

Pa   RQVSVEVPGI-GRVSGDIAWGGNWFFLVAGH--GQRLAGDNLDALTAYTVAVQQALDD----QDIRGEDGGAIDH    206

Pa   IELFAD--DPHADSRNFVLCPGKAYDR SPCGT GTSAKLACLAADGKLLPGQPWRQASVIGSQFEGRYEWLDGQ    279


Pa   PGGPIVPTIRGRAHVSAEATLLLADDDPFAWGIRR- ----------------     314
```

Splice leader
acceptor sites

Signal Peptide

Polypyrimidine rich region

⇩

```
CCTTTTTCTTTTTAAAAACAAAAAAAATTCCGGGGGGAATATGGAACAGGGTATATGCGTAAAAGTGTCTGTCCCAAACAAAAATTTTT      90
                                               M  R  K  S  V  C  P  K  Q  K  F  F             12
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTTGTGTGTGTTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTTGACCAAAAATAT     180
 F  S  A  F  P  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  D  Q  K  Y         42
AAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAAGGGAAATTATGCGATTT     270
 K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  E  H  Q  Q  K  R  E  I  M  R  F        72
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT     360
 K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N       102
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT     430
 M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F       132
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT     520
 G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H       162
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA     610
 N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T       192
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATTT     700
 P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F       222
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT     790
 L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I       252
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT     880
 V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N       282
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCA     970
 R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A       312
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT     1060
 N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y       342
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCCGA     1150
 A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R       372
ATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG     1240
 I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M       402
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT     1330
 G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  ·                               423
GGGAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTATTTTTTATTTTTATTAGTGCATT     1420
ATTATTAAATTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGCTTTTAT     1510
TTTCATTTTCATCTTCCTACCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGGTCTTGTGGGTGGAGGAGTCCTAAATCCCGCACCTCGG     1600
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT                                                 1651
```

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(d)   Nucleotide sequence and peptide sequence TcPA45

```
                                                M  R  K  S  V  C  P  K  Q  K  F
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTTGTGTGTTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTTGACCAAAATAT    190
 F  S  A  F  P  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  D  Q  K  Y         42
AAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAAGGGAAATTATGCGATTT    270
 K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  E  H  Q  Q  K  R  E  I  M  R  F         72
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT    360
 K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N        102
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT    450
 M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F        132
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT    540
 G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H        162
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA    630
 N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T        192
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATTT    720
 P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F        222
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT    790
 L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I        252
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT    880
 V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N        282
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCA    970
 R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A        312
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT    1060
 N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y        342
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA    1150
 A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R        372
ATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG    1240
 I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M        402
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT    1330
 G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  .                                 423
GGGAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTATTTTTATTTTTATTAGTGCATT     1420
ATTATTAAATTTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGCTTTTAT    1510
TTCATTTTCATCTTCCTACCCAACCCCCTTGGTTCCACCGGTCGCCGGCGGGGTCTTGTGGGTGGAGGAGTCCTAAATCCCGCACCTCGG    1600
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT                                                 1651
```

┌─────────────────────────────┐
│ Polyadenilation site   │
└─────────────────────────────┘

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

┌──┐
│   (a)   Nucleotide sequence and peptide sequence TcPA45   │
└──┘

```
                                    CGAACAGGGCAGGAAAAGCTTCTGTTTGACCAA AAATAT  270
 F  S  A  F  P  F  F  F  F  F  C  V  F  P  L  I  S  R  T  G  Q  E  K  L  L  F  D  Q  K  Y   72
AAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACAAAAAAGGGAAATTATGCGATTT  360
  K  I  I  K  G  E  K  K  E  K  K  K  N  Q  R  A  N  R  R  R  E  H  Q  Q  K  R  E  I  M  R  F  102
AAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTCGAAT  430
  K  K  S  F  T  C  I  D  M  H  T  E  G  E  A  A  R  I  V  T  S  G  L  P  H  I  P  G  S  N   132
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGTTT  520
  M  A  E  K  K  A  Y  L  Q  E  N  M  D  Y  L  R  R  G  I  M  L  E  P  R  G  H  D  D  M  F   162
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT  610
  G  A  F  L  F  D  P  I  E  E  G  A  D  L  G  M  V  F  M  D  T  G  G  Y  L  N  M  C  G  H   192
AACTCAATTGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACA  700
  N  S  I  A  A  V  T  A  A  V  E  T  G  I  V  S  V  P  A  K  A  T  N  V  P  V  V  L  D  T   222
CCTGCGGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATTT  790
  P  A  G  L  V  R  G  T  A  H  L  Q  S  G  T  E  S  E  V  S  N  A  S  I  I  N  V  P  S  F   252
TTGTATCAGCAGGATGTGGTGGTTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT  380
  L  Y  Q  Q  D  V  V  V  V  L  P  K  P  Y  G  E  V  R  V  D  I  A  F  G  G  N  F  F  A  I   282
GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT  970
  V  P  A  E  Q  L  G  I  D  I  S  V  Q  N  L  S  R  L  Q  E  A  G  E  L  L  R  T  E  I  N   312
CGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTGGACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCA  1060
  R  S  V  K  V  Q  H  P  Q  L  P  H  I  N  T  V  D  C  V  E  I  Y  G  P  P  T  N  P  E  A   342
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCAGCGCCAAGATGGCAACACTTTAT  1150
  N  Y  K  N  V  V  I  F  G  N  R  Q  A  D  R  S  P  C  G  T  G  T  S  A  K  M  A  T  L  Y   372
GCCAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCA  1240
  A  K  G  Q  L  R  I  G  E  T  F  V  Y  E  S  I  L  G  S  L  F  Q  G  R  V  L  G  E  E  R   402
ATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTTATCATG  1330
  I  P  G  V  K  V  P  V  T  K  D  A  E  E  G  M  L  V  V  T  A  E  I  T  G  K  A  F  I  M   423
GGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTATT  1420
  G  F  N  T  M  L  F  D  P  T  D  P  F  K  N  G  F  T  L  K  Q  *
GGGAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTTATTTTTATTTTTTATTTTTATTAGTGCATT  1510
ATTATTAAATTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGCTTTTAT  1600
TTTCATTTTCATCTTCCTACCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGGTCTTGTGGGTGGAGGAGTCCTAAATCCCGCACCTCGG  1651
AGGAATAAACATATTTCAATTTCATATCTTGGAATCAAAAGGCAT
```

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(b)   Nucleotide sequence and peptide sequence TcPA45

Signal peptide

M R K S V C P K Q K F F

Nucleotide sequence of signal sequence TcPA45

ATG CGA TTT <

K R F

SEQ ID NO : 12

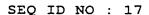5'TTICCRAADATIACIACGTT 3'

SEQ ID NO : 13

5' ATHGCITTYGGIGGIAAYTTT 3'

SEQ ID NO : 14

5' TTICCRAADATIACIACGTT 3'

SEQ ID NO : 15

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'

SEQ ID NO : 16

5' CTGAGCTCGACCAGATCTATCTGC 3'

SEQ ID NO : 17


```
   1 cctttttctt tttaaaaaca aaaaaaattc cgggggggaat atggaacagg gtatatgcgt
  61 aaaagtgtct gtcccaaaca aaaatttttt ttttccgcct tcccattttt ttttttttt
 121 tgtgtgtttc ccttgatctc tcgaacaggg caggaaaagc ttctgtttga ccaaaaatat
 181 aaaattatta agggcgagaa aaaagaaaag aaaaaaaatc aacgagcaaa caggagagaa
 241 caccaacaaa aaagggaaat tatgcgattt aagaaatcat tcacatgcat cgacatgcat
 301 acggaaggtg aagcagcacg gattgtgacg agtggtttgc cacacattcc aggttcgaat
 361 atggcggaga agaaagcata cctgcaggaa aacatggatt atttgaggcg tggcataatg
 421 ctggaaccac gtggtcatga tgatatgttt ggagcctttt tatttgaccc tattgaagaa
 481 ggcgctgact tgggcatggt attcatggat accggtggct atttaaatat gtgtggacat
 541 aactcaattg cagcggttac ggcggcagtt gaaacgggaa ttgtgagcgt gccggcgaag
 601 gcaacaaatg ttccggttgt cctggacaca cctgcggggt tggtgcgcgg tacggcacac
 661 cttcagagtg gtactgagag tgaggtgtca aatgcgagta ttatcaatgt accctcattt
 721 ttgtatcagc aggatgtggt ggttgtgttg ccaaagcccт atggtgaagt acgggttgat
 731 attgcatttg gaggcaattt tttcgccatt gtcccgcggg agcagttggg aattgatatc
 841 tccgttcaaa acctctccag gctgcaggag caggagaac ttctgcgtac tgaaatcaat
 901 cgcagtgtga aggttcagca ccctcagctg ccccatatta acactgtgga ctgtgttgag
 961 atatacggtc cgccaacgaa cccggaggca aactacaaga acgttgtgat atttggcaat
1021 cgccaggcgg atcgctctcc atgtgggaca ggcaccagcg ccaagatggc aacactttat
1081 gccaaaggcc agcttcgcat cggagagact tttgtgtacg agagcatact cggctcactc
1141 ttccagggca gggtacttgg ggaggagcga ataccggggg tgaaggtgcc ggtgaccaaa
1201 gatgccgagg aagggatgct cgttgtaacg gcagaaatta ctggaaaggc tttatcatg
1261 ggtttcaaca ccatgctgtt tgacccaacg gatccgttta agaacggatt cacattaaag
1321 cagtagatct ggtagagcac agaaactatt ggggaacacg tgcgaacagg tgctgctacg
1381 tgaagggtat tgaatgaatc gtttttttt atttttattt tttatttta ttagtgcatt
1441 attattaaat ttttttttg tttggggtt tcaacggtac cgcgttggga gcaggaagc
1501 gatagcggcc ggacaatttt ttgcttttat tttcattttc atcttcctac ccaacccct
1561 tggttccacc ggtcgcggcg gggtcttgtg ggtggaggag tcctaaatcc cgcacctcgg
1621 aggaataaac atatttcaat ttcatatctt ggaatcaaaa ggcat
```


98

SEQ ID NO : 18

WIIK

SEQ ID NO : 19

IVTGSLPDISG

SEQ ID NO : 20

ATNVPVVLDTPAGLVR

SEQ ID NO : 21

VDIAFGGNF

SEQ ID NO : 22

NVVIFGNR

SEQ ID NO : 23

MATLYAK

SEQ ID NO : 24

5' TCCGTATCCATGTCGATGC 3'

SEQ ID NO : 25

5' TATTATTGATACAGTTTCTG 3'

SEQ ID NO : 26

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'